Supporting Information

Intermediate in the O-O Bond Cleavage Reaction of an Extradiol Dioxygenase

Elena G. Kovaleva and John D. Lipscomb*

Department of Biochemistry, Molecular Biology, and Biophysics,
University of Minnesota, Minneapolis, MN 55455

*To whom correspondence should be addressed. E-mail: lipsc001@umn.edu

EXPERIMENTAL PROCEDURES

Materials. All chemicals for cell growth and enzyme purification were obtained from either Sigma-Aldrich Chemical Company or Fisher Scientific and used without purification. 4-Sulfonylcatechol was the generous gift of James W. Whittaker.

Site-directed mutagenesis, expression and purification. The E323L mutation of homoprotocatechuate 2,3-dioxygenase (HPCD) from *Brevibacterium fuscum* (EC 1.13.11.15) was created in the plasmid pYZW204 (*1*) using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The primers used to create the L323E mutation (forward, 5'-AGGAGATCATC**TTA**CGCACCGACGA-3', reverse 5'-TCGTCGGTGCG**TAA**GATGATCTCCT-3', site of mutation shown in bold) were synthesized and the desired mutation was verified by gene sequence analysis in the Microchemical Facility at the University of Minnesota. The sequencing primer designed for verification of desired mutation is 5'-GCACATGCGCTACGACCTGTACTC-3'. The E323L mutant of HPCD was expressed and purified in *E. coli* as described previously (*2*). ESI MS analysis showed that the mass of the mutated enzyme decreased by 16 mass units as expected. The kinetic parameters for the wild-type and E323L variant are shown in Table S1.

Table S1. Steady-state kinetic parameters for HPCD and E323L-HPCD at pH 7.5 [a]

Substrate	Homoprotocatechuate		4-Sulfonyl catechol	
HPCD Enzyme	Wild-type	E323L	Wild-type	E323L
V, s^{-1}	11.2 ± 0.7	9.87 ± 0.08	1.05 ± 0.05	1.32 ± 0.04
V/K_m, $mM^{-1}s^{-1}$	580 ± 40	580 ± 20	3.6 ± 0.3	3.5 ± 0.4
K_m, μM	19.4 ± 0.8	17.1 ± 0.7	290 ± 40	370 ± 50

[a] Determined from the fit of the initial (HPCA) and steady-state (4SC) velocity values at varied substrate concentrations to the Michaelis-Menten equation. Product formation was monitored by an increase in absorbance at 380 nm in air-saturated 50 mM Tris-HCl (pH 7.5) buffer using molar extinction coefficients for ring-cleavage products of HPCA (36 $mM^{-1}cm^{-1}$) and 4SC (20.4 $mM^{-1}cm^{-1}$) .

Crystallization and preparation of complexes. The crystals of E323L HPCD were grown in 18% PEG 8000, 0.1 M calcium acetate, 0.1 M sodium cacodylate, pH 6.5 or 7.0 (Hampton Research) by a hanging

or sitting drop method at 19 °C either inside an anaerobic glove box (Belle Technology) or under air. 20% glycerol was used as the cryoprotectant. General procedures for complex preparation and cryofreezing were as described previously (*2*). Reaction was initiated in the glove box by transfer of E323L HPCD crystal into equilibrated mother liquor solution containing 10 mM 4-sulfonylcatechol (4SC) at pH 7.0. After soaking for 30-60 min under approximately 5 – 10 ppm oxygen inside the glove box, crystals were flash frozen in liquid nitrogen before removing them from the glove box.

Data collection and refinement. X-ray data was collected at 100 K and 0.9790 Å at the Structural Biology Center (beamline 19BM), Advanced Photon Source, Argonne National Laboratory. Diffraction data were processed with HKL2000 (*3*). The structures were refined by cycles of restrained refinement with Refmac5.2.0019 (*4*) as part of the CCP4 program suite (*5*) and model building using Coot (*6*). The coordinates for full length wild-type HPCD (PDB code 2IG9(*2*)) were used as a molecular replacement model. Link restraints to Fe were removed from the refinement to avoid bias in the refined Fe-ligand distances. NCS restraints were not used during refinement, and the 4 subunits of the single enzyme molecule present in the asymmetric unit were refined independently. X-ray data processing and refinement statistics are summarized in Table S2.

Table S2. X-Ray data collection and refinement statistics[a].

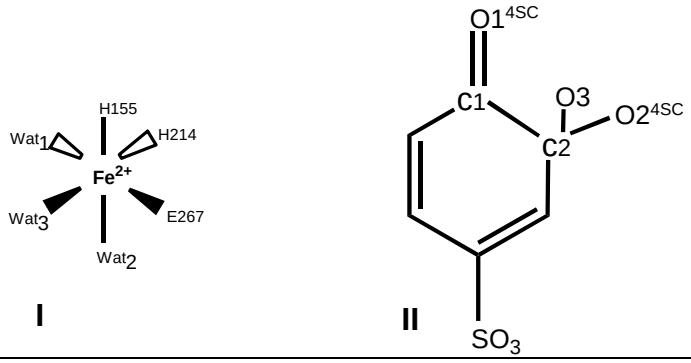
Dataset	E323L HPCD	E323L HPCD reacted with 4SC/O_2
PDB code	3ECJ	3ECK
Spacegroup	$P2_12_12$	$P2_12_12$
Monomer/AU	4	4
Cell dimensions (Å)	110.71, 163.40, 101.58	110.74, 163.18, 101.37
Cell angles (deg)	90, 90, 90	90, 90, 90
Resolution range[a] (Å)	50-1.65 (1.71)	55 – 1.60 (1.66)
Reflections (observed/unique)	952580/214697	907740/222399
R_{merge}[a, b] (%)	0.061 (0.469)	0.052 (0.342)
Mean $\langle I\rangle/\sigma\langle I\rangle$ [a]	18.73 (1.65)	19.22 (3.07)
Completeness (%)[a]	97.2 (96.5)	92.4 (73.2)
R, R_{free}, test (%)[c]	16.9, 19.4, 5	17.4, 19.8, 5
RMSD[d] bond length (Å)	0.012	0.011
RMSD[d] angles (deg)	1.302	1.297
ESU[e] (Å)	0.058	0.054
Ramachandran Plot		
Allowed regions (%)	99.7	99.7
Additional regions (%)	0.3	0.3
Disallowed regions (%)	0.0	0.0

[a] Values for the highest resolution shell are given in parenthesis. [b] $R_{merge} = \sum|I_i - \langle I\rangle|/\sum|I_i|$, where I_i is the integrated intensity of a given reflection, and $\langle I\rangle$ is the mean value for that reflection. [c] $R = (\sum|F_{obs} - kF_{calc}|)/\sum|F_{obs}|$, where k is a scale factor. The R_{free} value was calculated with the indicated percentage of reflections not used in the refinement. [d] Root-mean-square deviation (RMSD) from ideal geometry in the final models. [e] Estimated overall coordinate error (ESU) based on maximum likelihood.

Ligand refinement and molecular modeling. After refinement of the protein and solvent was completed, 4SC was used as the initial ligand model in the active sites. Based on the resultant F_{obs}-F_{calc} map, appropriate ligands were modeled into the density to obtain the best fit of the data (see additional information below). Initial ligand dictionaries for the 4SC-gem diol adduct were produced with Monomer Sketcher (*5*, *7*).

Model bias associated with ligand refinement was evaluated by examining ligand-omit difference maps. Active site ligands were removed from the final refined protein model, and ligand-free model was allowed to refine with 5 cycles of restrained refinement. The resultant positive density in the difference (ligand omit) F_{obs}-F_{calc} maps was found to agree with the refined positions and structures of the active site ligands in the full models (e.g. Figure 1A, Figure S1A and B). Interatomic distances and angles are summarized in Table S3.

Table S3. Summary of distances and angles for ligand coordination to Fe[a]



	Ligands coordinated to Fe in resting form and 4SC gem diol adduct of E323L HPCD			
	E323L (PDB 3ECJ)	E323L reacted with 4SC and O_2 (PDB 3ECK)		
Distances/Angles	Solvent[b]	Solvent[c]	4SC gem diol adduct[d]	4SC gem diol adduct[e]
Fe – Wat1	2.15 Å	2.20 Å	2.26 Å	2.29 Å
Fe – Wat2 [f]	2.27 Å	2.28 Å		
Fe – Wat3	2.24 Å	2.24 Å		
Fe – $O1^{4SC}$			2.16 Å	2.17 Å
Fe – $O2^{4SC}$			2.20 Å	2.19 Å
Fe – $O3^{4SC}$			3.61 Å	3.50 Å
Wat1 – $O3^{4SC}$			2.68 Å	2.59 Å
∠ O3 – C2 – $O2^{4SC}$			104.2 °	103.4 °
∠ O3 – C2 – C1			119.4 °	114.0 °
∠ O3 – C2 – C3			93.5 °	97.7 °
∠ $O2^{4SC}$ – C2 – C3			112.2 °	112.6 °
∠ $O2^{4SC}$ – C2 – C1			114.0 °	116.6 °

[a] For simplicity, nomenclature based on the molecular origin of the atom is used. Structures shown represent solvent bound to Fe in the free enzyme (I) and gem-diol adduct of 4SC (II). [b] Average distances calculated for 4 monomeric subunits of the free form of E323L enzyme (PDB code 3ECJ). Errors for distances were 2% or less. [c] Average distances calculated for subunits A and B in the structure of E323L enzyme reacted with 4SC and oxygen (PDB code 3ECK). Errors for distances 1 % or less. [d] Values corresponding to structure shown in Figure 1A or S1B (subunit C). [e] Values corresponding to structure shown in Figure S1A (subunit D). [f] In the high-resolution structures of HPCD, slightly asymmetric rather than spherical electron density is apparent at one of the solvent coordination sites. No small ligand that can account for the observed density is present in purification or crystallization solutions.

Site Occupancy. Based on the average temperature factors, ligand occupancies for gem diol adduct were estimated to be ~85 % and 65 % in subunits C and D, respectively. The ligand in subunit D was present at lower occupancy, but the same gem-diol features as those present in subunit C (6-membered ring, sulfonyl and gem-diol moieties) were observed as shown in Figure S1. This observation, in combination with the similarity in molecular geometry of the structures in C and D, suggests that subunit D is either occupied by the gem diol intermediate or vacant. Subunits A and B were found not to have 4SC ligand bound in any form, and were refined as a resting state with Fe-bound solvent. The structures of the putative gem-diol intermediate do not have atoms with anomalously large or small B-factors, and similar structures are determined from data collected in house and at the synchrotron, suggesting that the new atom associated with substrate C2 is not introduced by a reaction associated with X-ray irradiation.

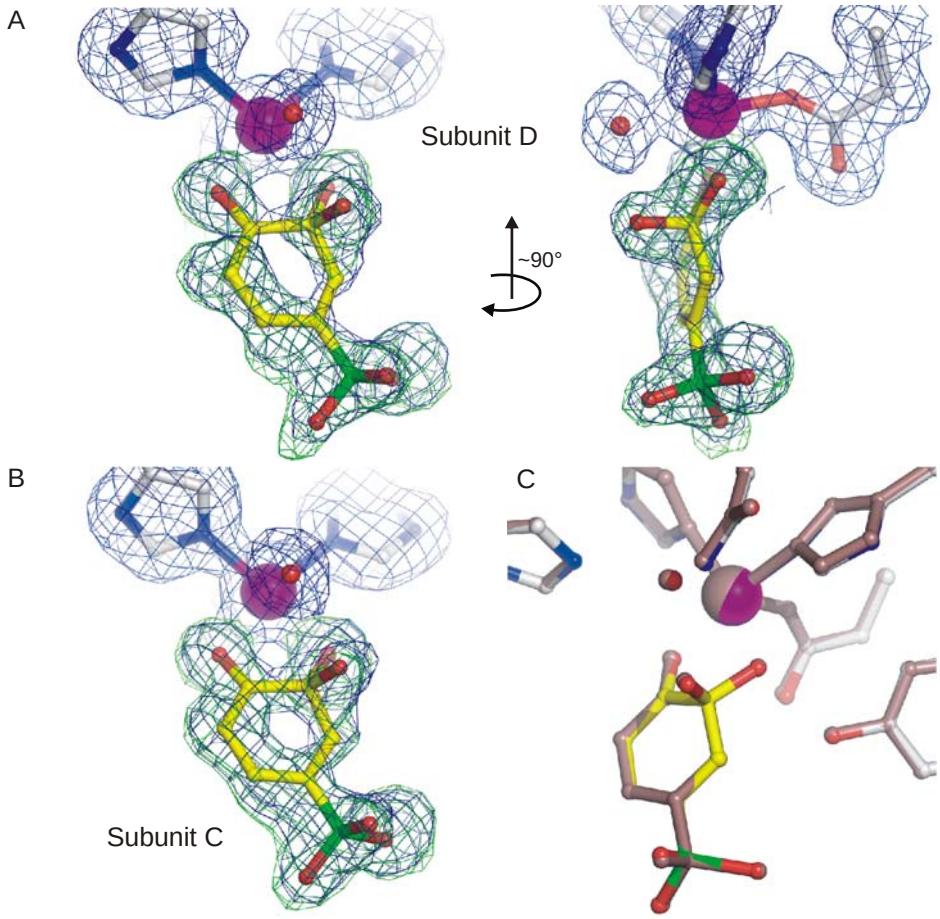


FIGURE S1: The same intermediate is observed in subunits C and D of the E323L mutant of HPCD reacted *in crystallo* with 4SC (PDB code 3ECK). (A) Two views of the structure of the putative gem-diol intermediate in the subunit D where it is present at slightly lower occupancy than in the subunit C. (B) View of the intermediate in the subunit C for comparison with panel A (also see Figure 1). The blue $2F_{obs}$-F_{calc} maps are contoured at 0.8 σ (A) and 1.0 σ (B). The green ligand-omit difference maps (F_{obs}-F_{calc}) were computed without ligands present in the model and are contoured at +3.0 σ (A), +3.5 σ (B). Atom color code: gray, carbon (enzyme residues); yellow, carbon (ligands); blue, nitrogen; red, oxygen; green, sulfur; purple, iron. (C) Overlay of the active site structures in the subunits C and D. Active site residues and ligand in subunit C are color coded as described in (B); corresponding model of subunit D is colored in bronze.

Halide site. As described previously, strong spherical density is observed in the pocket created by His248/Arg243/Arg293 at the bottom of the substrate-binding cleft in the resting form, and it is assigned to Cl^- anion (*2*). The position of the bound anion (Cl^-) in the free enzyme essentially overlaps with the position of the oxygen atom from the sulfonyl substituent of the 4SC-derived ligands. Since ligands in the

reactive complex structures appear to be present at incomplete occupancy, it is likely that the anion would occupy the His248/Arg243/Arg293 pocket in their absence. Inclusion of Cl^- anion at partial occupancy during refinement normalized the temperature factor for the oxygen atoms of sulfonyl moiety and eliminated slight excess of positive F_{obs}-F_{calc} density at the site.

Reproducibility. Data sets were collected for E323L crystals incubated with 4SC under low O_2 for various periods of time, under different buffer conditions, using both an in-house and a synchrotron source. In total, 5 data sets showed evidence for the gem diol intermediate structure in subunits C and D. The similarly studied WT HPCD crystals also revealed the gem diol species under some conditions, but not others. In cases where the gem diol was observed, it was not as well resolved as for the E323L variant, and thus we presume that more than one species was present.

ADDITIONAL SUPPORTING INFORMATION

Ligand modeling. Active site ligands were not included in the initial refinement of the protein models and were built into the density from F_{obs}-F_{calc} map. Several 4SC-derived intermediate species drawn from the mechanistic hypotheses that we and others have put forward were considered during sequential fitting of the density features observed in the active sites of subunits C and D. These included 4SC, 4SC-derived intermediates (semiquinone, alkylperoxo, epoxide and lactone) as well as open-ring product.

The ligand omit difference density (Figures 1A and S1A and B) clearly illustrates that none of the active sites contain 4SC in an unaltered form. It is also clear that 4SC-derived ligand in subunits C and D must have a non-planar, 6-membered ring and a well-defined position of the sulfonyl moiety. These features eliminate the 7-membered ring lactone and open-ring products as possible ligands. C-2 of the ligand ring exhibits tetrahedral character, and extra electron density at that position is well defined and separated from the solvent bound to the Fe. As such, neither semiquinone nor alkylperoxo intermediate species can account fully for the observed density at C2. The epoxide intermediate species with a single additional atom between C2 and C3, was also used to fit the data. However, refinement of the epoxide model resulted in significant lengthening and effectively breaking of the O3-C3 bond (2.3 Å), producing a refined model which is consistent with geometry of a gem-diol species rather than an epoxide.

As seen for the complexes with intact ring derivatives of 4-nitrocatechol (*2*), the $2F_{obs}$-F_{calc} density for C3-C4 bond of 4SC gem diol adducts was also found to be weaker than that for other C-C bonds of the ring. The weaker density for C3-C4 bond of 4SC/4NC adducts may be due to ring puckering or partial single bond character during species interconversion or proximity to the anionic moiety of the ring substituent stabilized by the His248/Arg243/Arg293 pocket. Alternatively, it may be due to dynamic fluctuations resulting from non-optimal binding due to shorter ring substituent (nitro or sulfonyl group vs carboxymethyl).

Effect of the E323L mutation and the 4SC gem diol on HPCD structure. As illustrated in Figure S2, residue 323 is located on the surface of the enzyme. The superposition analysis of the wild-type (Glu323) and E323L enzymes indicates that the structures of the individual subunits as well as overall tetrameric assembly are unaltered by this mutation as indicated by rmsd values of 0.35-0.45 Å for all atoms or 0.12-0.14 Å for main chain atoms only. The E323L mutation does, however, have an affect on the crystal packing interactions as indicated by the increase in cell dimensions (Figure S2) and changes at the symmetry interface (Figure S3).

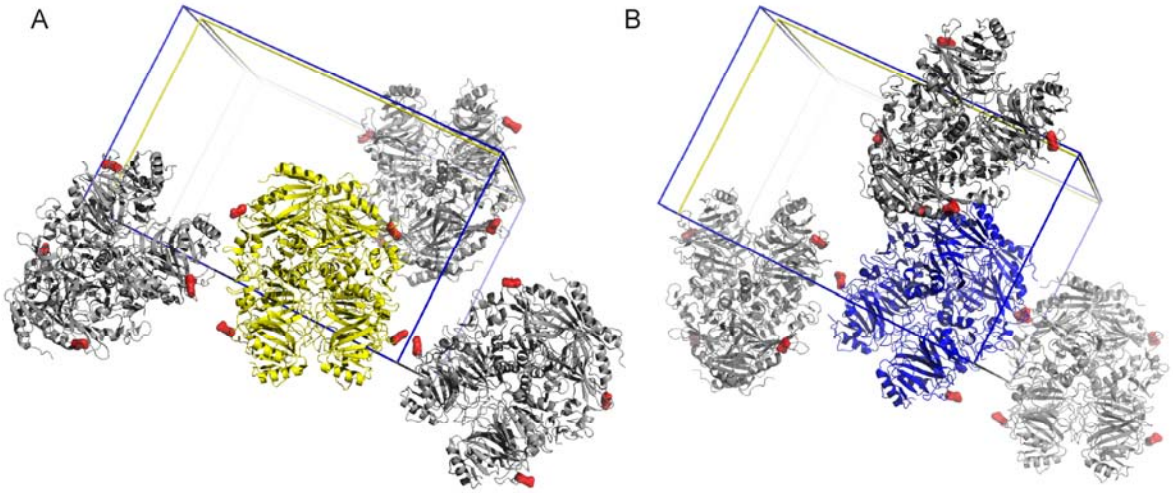


FIGURE S2: Structure of the (A) wild-type and (B) E323L HPCD illustrating location of residue 323 (red), relative orientations of symmetry related tetramers that have interactions in the residue 323 region, and changes in unit cell dimensions due to the single E323L mutation. Color code: yellow, HPCD tetramer (PDB code 2IG9) and unit cell; blue, E323L tetramer (PDB code 3ECJ) and unit cell; gray, symmetry-related molecules; red, with Glu323 (panel A) or Leu323 (panel B)..

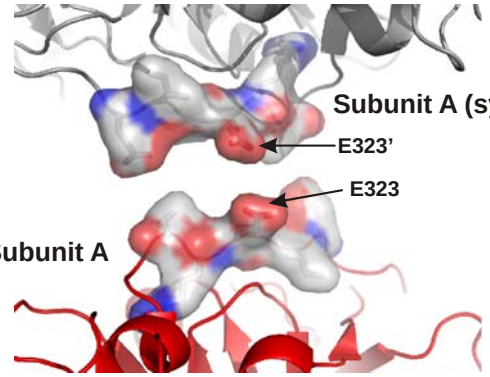
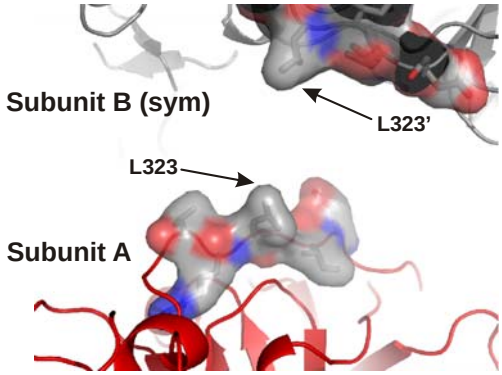
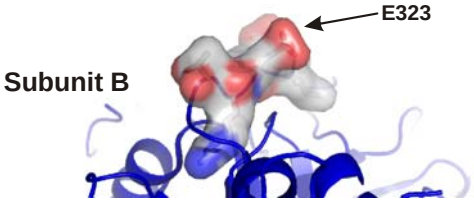
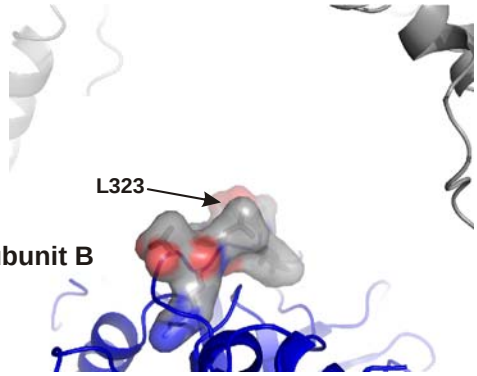
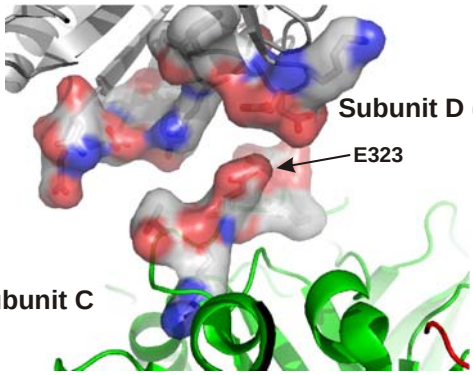
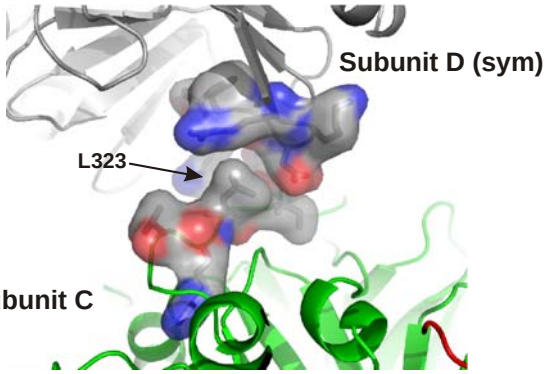
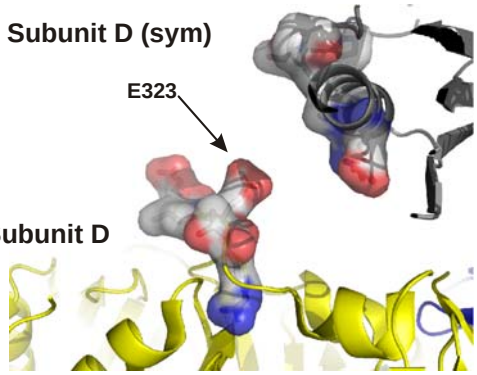
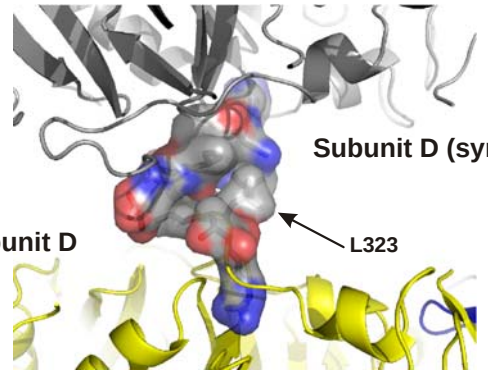
Wildtype	E323L Mutant
	
	
	
	

FIGURE S3: Representative changes in crystal packing interactions due to E323L mutation. Color code for enzyme subunits represented as cartoon: red, subunit A; blue, subunit B; green, subunit C; yellow, subunit D; gray, subunits from different symmetry-related molecules. Color code for residue side-chains at the symmetry interface near residue 323 shown as stick and surface model: gray, carbon; red, oxygen; blue, nitrogen. (Left column) Symmetry interface region near residues 320-325 for subunits of the wild-type enzyme (PDB code 2IG9). Residues from the symmetry-related molecules interacting with subunit A are 320-325 (subunit A (sym)), with subunit C are 45-49, 105-110 (subunit D (sym)), with subunit D are 8-11, 91,92,95,96 (subunit D (sym)). (Right column) Symmetry interface region near residues 320-325 for subunits of the E323L enzyme (PDB code 3ECJ). Residues from the symmetry-related molecules interacting with subunit A are 320-325 (subunit B (sym)), with subunit C are 104-111 (subunit D (sym)), with subunit D are 207-209, 320-325 (subunit D (sym)).

REFERENCES

1. Wang, Y. Z., and Lipscomb, J. D. (1997) Cloning, overexpression, and mutagenesis of the gene for homoprotocatechuate 2,3-dioxygenase from Brevibacterium fuscum, *Protein Exp. Purif. 10*, 1-9.
2. Kovaleva, E. G., and Lipscomb, J. D. (2007) Crystal structures of Fe^{2+} dioxygenase superoxo, alkylperoxo, and bound product intermediates, *Science 316*, 453-457.
3. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode in *Macromolecular Crystallography, part A* (Carter, C. W., Jr. , and Sweet, R. M., Eds.), pp 307-326, Academic Press, New York.
4. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240-255.
5. CCP4. (1994) *Acta Crystallogr. D50*, 760-763.
6. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. E60*, 2126-2132.
7. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. J. (2003) A graphical user interface to the CCP4 program suite, *Acta. Cryst. D59*, 1131-1137.